COMPLEX VISUAL INFORMATION PROCESSING: A TEST FOR PREDICTING NAVY PRIMARY FLIGHT TRAINING SUCCESS

Tommy R. Morrison

Naval Aerospace Medical Research Laboratory
Naval Air Station
Pensacola, Florida 32508-5700

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) NAMRL-1338	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Naval Aerospace Medical Research Laboratory 6b. OFFICE SYMBOL (If applicable) 032	7a. NAME OF MONITORING ORGANIZATION Naval Air Systems Command
6c. ADDRESS (City, State, and ZIP Code) NAS, Pensacola, FL 32508-5700	7b. ADDRESS (City, State, and ZIP Code) Washington, DC, 20361-3300
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Office of Naval Research 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code) 800 North Quincy Street Arlington, VA 22217-5000	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 61153N PROJECT NO. WR04210 TASK NO. 01 WORK UNIT ACCESSION NO. 6142

11. TITLE (Include Security Classification)
(U) Complex Visual Information Processing: A Test for Predicting Navy Primary Flight Training Success

12. PERSONAL AUTHOR(S)
Tommy R. Morrison

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Interim	FROM 1986 TO 1987	June 1988	22

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Pilot Selection Test Validation, Predicting Aviator Performance, Cognitive Processes in Aviation

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

A continuing problem in naval aviation is how to identify, prior to training, those individuals having the greatest likelihood to succeed in flight training. Since World War II the US Navy has been concerned with the development of selection tests to improve its capability to screen individuals desiring entry into flight training. Improved aviation personnel selection tests could reduce attrition, and consequently, the monetary losses associated with this attrition. Additionally, the Navy would expose fewer flight students to a potentially hazardous training situation. A research methodological problem associated with the development of candidate tests for predictive validation is: How do we design a test that measures abilities relevant for inclusion in validation studies?

A methodology, Random Sampling of Domain Variances, was developed as a product of the

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL J. O. HOUGHTON, CAPT MC USN Commanding Officer	22b. TELEPHONE (Include Area Code) (904) 452-3286 22c. OFFICE SYMBOL 00

DD Form 1473, JUN 86 Previous editions are obsolete. SECURITY CLASSIFICATION OF THIS PAGE

S/N 0102-LF-014-6603 UNCLASSIFIED

19. Abstract (Continued)

Augmentation of Human Factors Engineering Technology Efforts Program. This methodology provided the basis for developing a complex visual information task designed to require many basic cognitive processing demands thought to be required in performing aviation display tasks. Measures from this complex visual task were significant predictors against three criterion measures of US Navy Primary Flight Training success: Pass/Fail, Flight Grades, and Composite Scores. Number of correct responses to the visual task accounted for 7.5% of the predicted variance associated with the Pass/Fail criterion. A prediction model consisting of number of correct responses to this task, plus the Navy's presently used Biographical Inventory score accounted for 10.2% of the Pass/Fail criterion variance in this sample of 451 Student Naval Aviators. The present results indicate that this complex visual task may be measuring critical basic cognitive processes that mediate both general intelligence test performance and success in naval aviation.

This complex visual display task should be further evaluated in a cross-validation study and then considered for implementation as a performance-based selection test. The method of Random Sampling of Domain Variances should be considered as a methodological tool in the design of future tests to be selected for validation studies. (SDW)

SUMMARY PAGE

THE PROBLEM

Identifying individuals, before they are actually trained, who are most likely to succeed in flight training is a continuing problem in naval aviation. By improving aviation personnel selection tests, the U.S. Navy could reduce attrition and associated monetary losses and expose fewer flight students to potentially hazardous training situations. Since World War II, the Navy has funded efforts to improve selection tests, which screen individuals desiring entry into flight training. A methodological problem associated with the development of candidate tests for predictive validation is: How do we design a test that measures abilities relevant for inclusion in validation studies?

FINDINGS

Random Sampling of Domain Variances was a methodology developed as a product of the Augmentation of Human Factors Engineering Technology Efforts program. This methodology provided the basis for developing a complex visual information task, which was designed to require many fundamental cognitive processing demands that are presumably necessary in performing aviation display tasks. Measures from this complex visual task were significant predictors against three criterion measures of U.S. Navy primary flight training success: pass/fail, flight grades, and composite scores. The number of correct responses to the visual task accounted for 7.5% of the predicted variance associated with the pass/fail criterion. A prediction model of the number of correct responses to this task plus the Navy's current Biographical Inventory score accounted for 10.2% of the pass/fail criterion variance in a sample of 451 student naval aviators. The present results indicate that this complex visual task may be measuring critical basic cognitive processes that mediate both general intelligence test performance and success in naval aviation.

RECOMMENDATIONS

This complex visual display task should be further evaluated in a cross-validation study and then considered for implementation as a performance-based selection test. The method of Random Sampling of Domain Variances should be considered in the design of future tests selected for validation studies.

Acknowledgments

The author gratefully acknowledges the efforts of Dr. Robert J. Wherry, Jr., in developing the complex visual display task. Mr. Al Thomas and HM2 Solomon Eagles provided outstanding support in administering the test. Mr. Peter Collyer's computer programming support in implementing the test and performing statistical analyses are gratefully acknowledged. The editorial comments and suggestions provided by Ms. Kathy Mayer are appreciatively acknowledged.

INTRODUCTION

This effort resulted from a research program entitled "Augmentation of Human Factors Engineering Technology Efforts" (AHFETE), which was a 6-year Special Focus program funded by the Office of Naval Research, via the Naval Air Systems Command, to the Naval Aerospace Medical Research Laboratory. The program began in 1981 and successfully terminated in 1986. From 1983 through 1986, the Naval Air Development Center and the Naval Aerospace Medical Research Laboratory participated jointly in the AHFETE program.

The technical objective of the AHFETE program was "To develop techniques to quantify operator workload capacities in terms of systems demands in order to provide common quantitative units for assessing the effects of task, environmental, and operator variables on human effectiveness while performing tasks relevant to naval aviation" (1, p. 1). Due to the ever increasing array of instrumentation that aircrew must respond to and the concomitant increase in information processing demands, the need to consider human capabilities and limitations in systems design is critical (2,3). Improved assessment of these human capabilites in aviation-relevant tasks could provide performance-based criteria relevant to the design of emerging systems.

A major product of the AHFETE program was a methodology, Random Sampling of Domain Variances, to sample performance variance estimates that are generalizable back to the real-world display task domain (5,6). Other major products were the development and the application of hierarchical factor analytic techniques to identify specific perceptual processes and their processing times of data generated herein (7,8). Morrison (9) presented a less elaborate, basic multiple regression approach in which the regression weights associated with particular display task cognitive demands were interpreted as estimates of the times required for the processes to occur.

The primary purpose of this study was to investigate the utility of a complex visual task (CVT) in selecting naval aviators, who perform many of their required functions via various complex visual displays. The visual display task selected for the present study resulted from a previous function/task analysis of the P-3C Tactical Coordinator (4). A critical consideration was how to develop the experimental display task in a way to accurately sample the real-world display task domain.

The present investigation analyzed mean CVT performance scores. Performance scores pertaining to individual perceptual processes, which may be derived via the hierarchical factor analytic methods previously described (7,8), were not analyzed. If CVT measures appear useful in aviation selection, then the CVT test may be included in other selection test batteries currently being developed and validated at the Naval Aerospace Medical Research Laboratory (10,11). A secondary purpose was to examine the relationship between CVT measures and other predictors of Navy primary flight training success.

METHOD

SUBJECTS

Student naval aviators (441 males and 10 females) undergoing Navy primary flight training at Naval Air Station Whiting Field, Milton, FL, participated as subjects. All subjects had passed the Navy flight physical examination. Their age ranged from 20 to 30 yr old; mean age was 23.2 yr.

APPARATUS

The test station consisted of a booth in which the seated subject performed the complex visual task. A TV monitor and Caramate rear-projection slide system were positioned in front of the subject, with the monitor left of center and the projector right of center. The subject-to-screen viewing distance was about 51 cm. A response keypad with keys labeled 0 through 9, true, false, and enter was positioned directly in front of the subject. The keypad contained an Apple microcomputer, interfaced to a switching system, that controlled task presentation and recorded subject response times.

One hundred and twenty slides were presented on an illuminated display screen. The screen area was 15.25 cm x 15.25 cm, which was divided into quadrants by horizontal and vertical lines (each had a 1-mm stroke width). A 7.62-cm diameter circle (1-mm stroke width) was centered on the display screen. Objects presented on the display screen varied in shape (triangle, rectangle, pentagon); color (red, green, white); size (small = 8 mm L x 6 mm W, medium = 12 mm L x 8 mm W, large = 15 mm L x 10 mm W); heading (N, NE, E, SE, S, SW, W, NW); and screen location (random). The number of symbols per slide varied from 10 to 20. The dimensions of the symbol sizes were measured on the surface of the display screen. Figure 1 illustrates the display format and the three shapes and three sizes. Each symbol enclosed a solid black triangle, which indicated symbol heading and inferred "movement." In 60 slides, all symbols were the same color; in 13 and 47 slides, symbols were two or three colors, respectively. For all cases, each symbol was one color. Triangles represented airplanes; rectangles and pentagons represented aircraft carriers and destroyers, respectively. The three shapes were presented in all three sizes and all three colors.

The TV monitor presented the subject with questions written in capital letters 7 mm H x 5 mm W and stroke width approximately 1.5 mm. Target characteristics in questions to identify display symbols requiring the subject's response are given in Table 1. Questions differed by the amount and the type of information asked. A simple question was: "How many red carriers are on the screen?" This question required subjects to memorize and recall two types of target symbol information (color = red, shape = carrier) to successfully respond to the subsequently presented display screen. Although the question included the words "... on the screen," it did not specify a particular screen portion, hence, the subject did not have to remember where on the screen (e.g., upper, right, left, etc.) to search. A more difficult question was: "At least two small red destroyers heading south are in the upper screen portion (true or false)?" This question included six kinds of information to be memorized and recalled: 1) number of question objects = 2; 2) size = small; 3) color = red; 4) shape = destroyer, 5) heading = south; and 6) screen portion = upper.

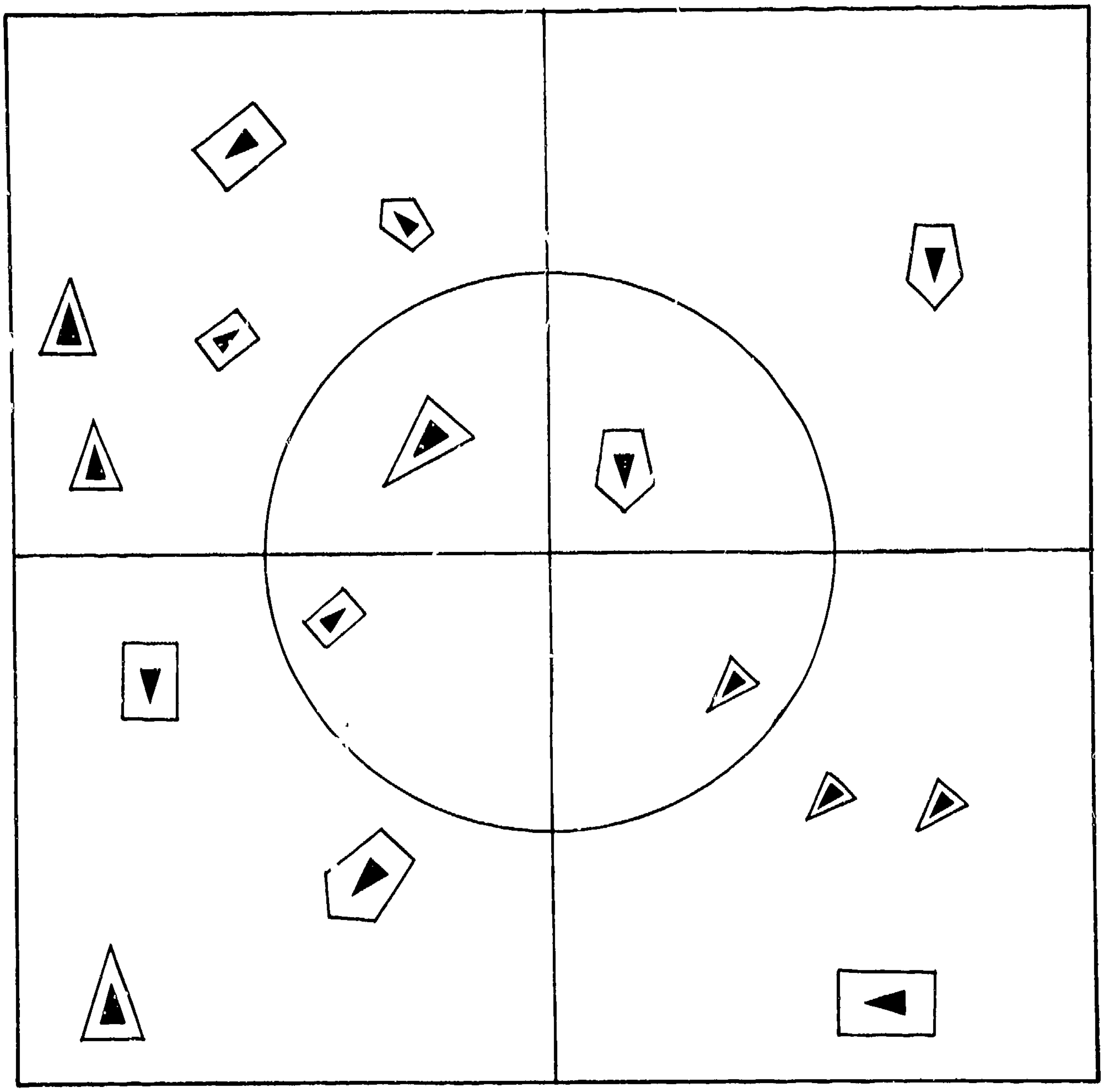

Figure 1. Typical display screen slide.

PROCEDURE

Taped verbal instructions (about 5 min) and programmed example slides were presented via a slide projector and display screen to each subject, who was seated in the test booth. The experiment followed the instructions and consisted of 3 slide group presentations, each containing 40 slides. The order of slide presentation within a slide group was constant, but the slide group order was random across subjects. The experimenter started the first trial of each slide group, and thereafter the experiment was self-paced. A trial consisted of the following: 1) A question appeared on the TV monitor; 2) the subject read the question; 3) the subject pressed "enter" (reaction time 1), which simultaneously removed the question from the TV monitor and presented the display slide; 4) the subject visually examined the display slide and responded via the keypad (reaction time 2) in accordance with the immediately preceding question; 5) the display slide was removed from view; and 6) feedback was presented on the TV monitor. The next trial began when the subject pressed "enter" again. Reaction times 1 and 2 were measured in milliseconds. Subjects required 20-25 min to complete each slide group. The Statistical Analysis System (SAS) version 5.16 (12) was used to analyze the data.

TABLE 1. Target Characteristics of Questions Presented on the TV Monitor.

Target characteristic	Definition
Circle position	Target symbols were either inside or outside the circle.
Screen portion	Target symbols could be in the: upper, lower, right, left, upper-left, upper-right, lower-left, or lower-right position of the display screen. No specified screen area meant targets could be anywhere in the full screen.
"Movement" to/from circle	Target symbols could be "moving" to or away from the center of the circle, or would pass through or enter the circle. "Movement" was inferred.
Shape	Triangle = airplane; rectangle = carrier; pentagon = destroyer.
Color	Red, green, white.
Size	Small, medium, large.
Heading	North, East, South, West, NE, SE, SW, NW.
Each red object represented 2 objects	This required subject to count each red object as really representing two objects.
Assume all red objects are rotated 90 degrees to the right	This required subjects to mentally rotate certain display symbols.
Number of display objects	Total number of symbols on display screen.
Target symbols	Number of target symbols to search for on display screen.
Number of question objects	Number of target symbols specified in question.
Two shapes	Certain questions included two shapes, e.g., carriers and airplanes.
Two headings	Certain questions called for target objects heading in either of two specified directions.
Two sizes	Certain questions called for targets of two sizes.
Two colors	Certain questions called for targets that were either of two colors.

For purposes of comparison, currently used predictors and measures of flight training performance are provided in the following two sections.

PREDICTORS OF FLIGHT TRAINING SUCCESS

The following is a list and brief description of screening tests currently used by the U.S. Navy:

1. The Academic Qualification Test (AQT) is a 60-min test of general intelligence. The AQT-C is a converted AQT score: mean = 5; standard deviation = 1.96; range = 1-9 (i.e., AQT-Cs are stanine scores).

2. The Flight Aptitude Rating (FAR) is a composite score based on the Spatial Apperception Test, Mechanical Comprehension Test, and Biographical Inventory. The FAR-C is the converted stanine FAR.

3. The Spatial Apperception Test (SAT) is a 10-min test of the ability to orient in space or, specifically, to visualize the relationship between the attitude of an airplane and the territory it flies over. The SAT-C is a c·· erted half-step stanine score: mean = 10; standard deviation = 3.96 ge = 1-19.

4. Mechanical Comprehension Test (MCT) is a 40-min test of the ability to perceive physical relationships and handle familiar concepts of everyday mechanics. The MCT-C is a converted half-step stanine score.

5. The Biographical Inventory (BI) is an untimed questionnaire of personal history, expressions of interest and attitudes, and selected information items. The BI-C is a converted half-step stanine score.

6. The Officer Aptitude Rating (OAR) is derived by combining scores from the AQT and the MCT. The OAR-C is a converted score, i.e., a Navy standard score, where mean = 50; standard deviation = 10; range = 20-80. The AQT and the FAR are used to screen applicants for entry into naval flight training. The OAR is used to screen individuals in the process of becoming an officer in the U.S. Navy. If an individual has taken the OAR and subsequently seeks entry into naval flight training, then the AQT and the FAR are administered (even though the individual has already taken the AQT and the MCT as the OAR).

PRIMARY FLIGHT TRAINING MEASURES

The following is a list and brief description of grades/scores obtained by individuals undergoing Navy primary flight training:

1. Flight Grade (FG) is the overall average flight grade at the end of primary flight training. It is based on item grades from various training flights as follows: unsatisfactory = 1; below average = 2; average = 3; and above average = 4. Item grades from 37 aircraft flights and 16 training simulator "flights" are equally weighted and averaged to produce the FG.

2. Academic Grade (AG) is the overall academic grade achieved through primary flight training. The AG is a converted Navy standard score.

3. Composite Score (CS) is a combination of the FG, AI, and AG, which are weighted 70%, 15%, and 15%, respectively. Although the FG is the most important factor in determining pipeline assignment following primary flight training, the CS may be used as a "tie-breaker" for students who have obtained equal FGs.

4. Pass/Fail (P/F) is a dichotomous criterion measure pertaining to primary flight training outcome.

5. Aviation Indoctrination (AI) is the average of certain academic courses. For officers under instruction, the AI is an average of three courses: aerodynamics, engineering, and navigation. For aviation officer candidates, the AI is an average of the above three courses plus eight other courses, such as, officer leadership quality, U.S. seapower, military law, et cetera.

RESULTS

Tables 2 and 3 present the zero-order correlations between the predictors and criterion measures. For individuals who passed primary flight training, CS, AI, AG, and FG scores were available. Those scores were not available for individuals who failed, hence, they do not appear in Table 3. Similarly, no P/F measure appears in Table 2 (because everyone passed). The number of complex visual task (CVT) slide presentations responded to correctly was designated as CVT-NC. The CVT-RT1 was the time required to read the question; CVT-RT2 was the time to respond to the display. The CVT-RT1s and CVT-RT2s were computed from the correct responses.

Table 2 presents many significant zero-order correlations; all are in the expected direction. Age correlated negatively with the AQT, OAR, CS, AI, and FG, and CVT-NC. The number of prior flight hours (PFLT) was related positively to the BI and the CS. Individuals with more PFLT obtained higher BI scores, which was not surprising as the BI contains many aviation knowledge/interest items. The AQT was directly related to the FAR, SAT, MCT, OAR, CS, AI, AG, and CVT-NC and inversely related to age, CVT-RT1, and CVT-RT2. Subjects with higher AQT scores performed a greater number of CVT trials correctly. Since the CVT-RT1 is the time required to read and commit to short-term memory alphanumerically presented information, its inverse correlation with the AQT was expected. The correlations between the FAR and the SAT, MCT, and BI were expectedly high and positive, since the FAR is a composite of the latter three. Likewise, the FAR varied directly with the AQT, OAR, CS, AI, AG, FG, and CVT-NC and varied inversely with the CVT-RT2. The high correlation between the FAR and the OAR is expected, since, the MCT is a part of each. Subjects with better FAR scores responded faster to the CVT display slide and obtained higher overall CVT accuracy.

As shown in Table 2, the SAT varied directly with the MCT, OAR, CS, AI, FG, AQT, FAR, and CVT-NC; the SAT was inversely related to the CVT-RT1 and the CVT-RT2. Subjects with higher SAT scores read the CVT questions faster and responded faster to the CVT display, at a higher level of accuracy. The MCT varied directly with the BI, OAR, CS, AI, AG, FG, CVT-NC, AQT, FAR, and SAT. Higher MCT scores were associated with faster responses to the CVT display. The BI was directly related to the OAR, CS, AI, AG, FG,

PFLT, FAR, and MCT. The BI was not related to any CVT measures. The OAR varied directly with the CS, AI, AG, FG, CVT-NC, AQT, FAR, SAT, MCT, and BI. Subjects with higher OAR scores had faster CVT-RT2s and were younger.

The high positive correlations between the CS and the AI, AG, and FG were expected, as the CS is a composite of the latter three scores. Higher CS scores were associated with faster CVT-RT1 and CVT-RT2 response times, and a greater number of correct responses to the CVT. Additionally, the CS varied directly with the PFLT, AQT, FAR, SAT, MCT, BI, and OAR and varied inversely with age. The AI scores varied directly with the AG, FG, CVT-NC, AQT, FAR, SAT, MCT, BI, OAR, and CS and varied inversely with age. The AG varied directly with the FG, CVT-NC, AQT, FAR, MCT, BI, OAR, CS, and AI. The FG scores varied directly with the CVT-NC, FAR, SAT, MCT, BI, OAR, CS, AI, and AG and varied inversely with age, CVT-RT1, and CVT-RT2.

The CVT-NC score varied directly with the AG, FG, AQT, FAR, SAT, MCT, OAR, CS, and AI and varied inversely with age. The CVT-RT1 measure varied inversely with the AQT, SAT, CS, and FG and varied directly with the CVT-NC and the CVT-RT2. The CVT-RT2 measure varied inversely with the FG, FAR, AQT, SAT, MCT, OAR, and CS and varied directly with the CVT-RT1 and the CVT-NC. The general trend for the CVT-RT1 and the CVT-RT2 was that those who were faster reading the CVT question and responding to the CVT display also had higher predictor scores and obtained higher criterion scores. Greater accuracy on the CVT varied directly with longer CVT question reading times and longer CVT display response times. The high positive correlation between the CVT-RT1 and the CVT-RT2 indicates that subjects who read the CVT questions faster also responded to the display faster.

The pattern of zero-order correlations obtained in the P/F group presented in Table 3 was similar to that in Table 2. An important correlation in Table 3 was between the CVT-NC and the P/F. For data analysis purposes, a pass was coded as a "1" and a fail as "2," thus the negative correlation between the CVT-NC and the P/F indicates that successful completion of primary flight training was associated with higher accuracy on the CVT.

In summary, the important findings regarding performance on CVT and subsequent performance during primary flight training were: Response accuracy (CVT-NC) varied directly with the CS, AI, AG, FG, and P/F. Both the CVT-RT1 and the CVT-RT2 were inversely related to the CS and the FG.

TABLE 2. Correlation Matrix of Data from Student Naval Aviators Who Passed Primary Flight Training.*

	1	2	3	4	5	6	7	8	9	10
1. Age										
2. PFLT	.090									
3. AQT-C	-.174	-.070								
4. FAR-C	-.023	.022	.215							
5. SAT-C	-.026	-.088	.195	.590						
6. MCT-C	-.029	-.004	.310	.689	.240					
7. BI-C	-.030	.149	.053	.676	.074	.360				
8. OAR-C	-.137	-.046	.810	.537	.263	.771	.229			
9. CS	-.119	.099	.199	.230	.158	.249	.191	.289		
10. AI	-.103	.078	.359	.273	.107	.371	.184	.452	.654	
11. AG	-.082	.059	.325	.151	.068	.216	.135	.333	.713	.675
12. FG	-.098	.096	.093	.185	.141	.184	.156	.185	.927	.452
13. CVT-NC	-.110	.022	.337	.183	.211	.235	-.038	.360	.198	.240
14. CVT-RT1	-.006	.015	-.148	-.036	-.122	.035	-.007	-.089	-.104	-.067
15. CVT-RT2	.010	.032	-.099	-.151	-.123	-.154	-.056	-.168	-.128	-.044

	11	12	13	14
12. FG	.500			
13. CVT-NC	.233	.130		
14. CVT-RT1	-.067	-.098	.132	
15. CVT-RT2	-.013	-.132	.231	.500

*r = +/- .098; p < .05; df = 405.

TABLE 3. Correlation Matrix of Data from Student Naval Aviators including 406 Who Passed and 45 Who Failed.*

	1	2	3	4	5	6	7	8	9	10
1. Age										
2. PFLT	.106									
3. AQT-C	-.156	-.073								
4. FAR-C	-.005	.018	.230							
5. SAT-C	-.024	-.075	.197	.586						
6. MCT-C	-.005	-.033	.324	.686	.225					
7. BI-C	-.020	.156	.059	.683	.076	.354				
8. OAR-C	-.110	-.067	.812	.546	.254	.778	.234			
9. P/F	.080	.013	-.056	-.149	-.055	-.116	-.162	-.107		
10. CVT-NC	-.096	-.004	.327	.189	.189	.254	-.014	.366	-.274	
11. CVT-RT1	-.002	.002	-.146	-.013	-.114	.043	.023	-.082	-.058	.119
12. CVT-RT2	.016	.025	-.100	-.141	-.120	-.150	-.041	-.161	-.025	.216

	11
12. CVT-RT2	.501

*r = +/- .093; p < .05; df = 450.

Table 4 presents zero-order correlations between the AQT, MCT, SAT, and BI obtained in 1963 (13) and 1988 (the present study). The correlations between the AQT and the MCT, the AQT and the BI, the MCT and the SAT, and the SAT and the BI have remained quite stable over the past 25 years. Fisher's z coefficients were determined for the correlations between the AQT and the SAT, and the MCT and the BI, and the difference between the z coefficients for 1963 and 1988 was tested for significance (14). The difference of the correlations between the AQT and the SAT from 1963 to 1988 was significant, $p < .005$. Similarly, the difference of the correlations between the BI and the MCT from 1963 to 1988 was significant, $p < .0001$. Thus, during the past 25 years, the zero-order correlation between the MCT and the BI has increased while the correlation between the AQT and the SAT has decreased.

TABLE 4. Correlation Matrix of AQT, SAT, MCT, and BI Scores: 1963 and 1988.*

	1963 (N = 587)				1988 (N = 446)			
	AQT	MCT	SAT	BI	AQT	MCT	SAT	BI
AQT								
MCT	.362				.317			
SAT	.272	.250			.194	.223		
BI	.026	.179	.099		.060	.361	.078	

*Shoenberger et al. (13).

Separate stepwise multiple regression (R) analyses were performed with the three criterion measures: P/F, FG, and CS. The significance level for the F statistic for a variable to be added to the model was .05. After a variable was added, however, the stepwise method examined all the variables already included in the model and deleted any variable that did not produce an F statistic significant at .05. Table 5 presents the results of the regression analysis with P/F as the criterion. The predictor variables available for model entry were the MCT, BI, CVT-NC, and OAR; these obtained significant zero-order correlations with the P/F as shown in Table 3.

TABLE 5. Summary Results of Stepwise Multiple Regression Analysis with Pass/Fail as Criterion.

Step 2 BI-C entered.

R**2 = 0.10236
Adjusted R**2 = 0.09835
R = 0.31994

Source	df	Sum of squares	Mean square	F	p
Regression	2	4.06451709	2.03225854	25.54	0.0001
Error	448	35.64279999	0.07955982		
Total	450	39.70731707			

	Variable		Partial	Model		
Step	entered	removed	R**2	R**2	F	p
1	CVT-NC		0.0748	0.0748	36.33	0.0001
2	BI-C		0.0275	0.1024	13.73	0.0002

The first variable to enter, CVT-NC, accounted for 7.48% of the P/F variance; the BI entered second, accounting for 2.75%. The two-predictor model was highly significant, accounting for a total of 10.24% of the P/F variance (adjusted R**2 = .0984). Of all the predictors--age, PFLT, AQT, FAR, SAT, MCT BI, OAR, CVT-NC, CVT-RT1, and CVT-RT2--the two that accounted for significant amounts of the P/F variance were the BI and the CVT-NC. Note in Table 3 that no CVT measure attained significant zero-order correlations with the BI.

The results from the stepwise regression analysis with the FG as the criterion are presented in Table 6. The predictor variables that had obtained significant zero-order correlations (see Table 2) were available for model entry and were: age, SAT, MCT, BI, OAR, CVT-NC, CVT-RT1, and CVT-RT2. Since the FAR is a composite of the SAT, MCT, and BI, it was not available for entry.

TABLE 6. Summary Results of Stepwise Regression Analysis with Flight Grade as Criterion.

Step 5 OAR-C removed.

R**2 = 0.06825
Adjusted R**2 = 0.06130
R = 0.26126

Source	df	Sum of squares	Mean square	F	p
Regression	3	0.03507393	0.01169131	9.82	0.0001
Error	402	0.47879800	0.00119104		
Total	405	0.51387193			

Step	Variable entered	Variable removed	Partial R**2	Model R**2	F	p
1	OAR-C		0.0343	0.0343	14.34	0.0002
2	BI-C		0.0136	0.0479	5.78	0.0167
3	CVT-RT2		0.0101	0.0580	4.29	0.0390
4	CVT-NC		0.0151	0.0730	6.51	0.0111
5		OAR-C	0.0048	0.0683	2.07	0.1509

The predictors entered the regression model in the following order: OAR, BI, CVT-RT2, and CVT-NC. On the fifth step, the OAR was removed; the remaining three variables accounted for 6.83% of the FG variance. Because the OAR was removed from the model, another stepwise analysis was performed with only the BI, CVT-NC, and CVT-RT2 available for entry. The results from Table 7 show that the overall model was significant; R**2 accounted for 6.83% of the FG variance (adjusted R**2 = .0613). Of the 6.83% predicted variance, the BI accounted for 2.44%, CVT-NC accounted for 1.86%, and CVT-RT2 accounted for 2.53%.

TABLE 7. Summary Results of Stepwise Regression Analysis with Flight Grade as Criterion.

Step 3 CVT-RT2 entered.	
	$R**2$ = 0.06825
	Adjusted $R**2$ = 0.06130
	R = 0.26126

Source	df	Sum of squares	Mean square	F	p
Regression	3	0.03507393	0.01169131	9.82	0.0001
Error	402	0.47879800	0.00119104		
Total	405	0.51387193			

	Variable		Partial	Model		
Step	entered	removed	R**2	R**2	F	p
1	BI-C		0.0244	0.0244	10.09	0.0016
2	CVT-NC		0.0186	0.0430	7.83	0.0054
3	CVT-RT2		0.0253	0.0683	10.91	0.0010

Results of the stepwise regression analysis with the CS as the criterion are presented in Table 8. The predictors that obtained significant zero-order correlations (see Table 2) and consequently were available for model entry were: age, PFLT, AQT, SAT, MCT, BI, OAR, CVT-NC, CVT-RT1, and CVT-RT2. The FAR score was not available for entry because it is a composite of the SAT, BI, and SAT, which were available. Although the OAR is a composite of the AQT and the MCT, it is a separate administration, therefore, the OAR as well as the AQT and the MCT were available for entry. The predictors entered in the following order: OAR, BI, CVT-NC, and CVT-RT2. The OAR, BI, CVT-NC, and CVT-RT2 accounted for 8.35%, 1.64%, 1.40%, and 1.50%, respectively, of the obtained R**2 = .1289 (adjusted R**2 = .1202).

mation bits contained in the chunk. Presumably, individuals possessing superior capabilities to encode, maintain, and manipulate a greater amount of information in STM faster, might succeed at a higher rate in flight training. The results of the regression analyses presented above support this assertion.

Basic cognitive processing ability has been related to intelligence. Hunt et al. (21) listed the following basic cognitive processes, collectively referred to as "Current Information Processing" (CIP): processes by which people detect stimuli, locate them in time and space, integrate sequences of stimuli over time to form a stable percept, transform the percept during problem solving, coordinate them with respect to stored representations of highly overlearned codes, and abstract information from the percept and place it in long-term memory. Hunt et al. (21) compared subjects (college freshmen) having high-versus-low verbal ability (measured by the Washington Pre-College Test) and found that high verbals displayed the following superior CIP abilities: ability to make a rapid conversion from a physical representation to a conceptual meaning; ability to maintain information in short-term memory about the order of stimulus presentation; and ability to be more rapid in the manipulation of data in short-term memory. Hunt et al. (21) regarded information processing capacity as a more basic mental ability than the composite skill of verbal aptitude. Performance on verbal aptitude tests is largely determined by what a person knows and not just how rapidly someone can perform basic, simple, cognitive processes such as the CIPs described above.

The correlation between the CVT-NC and the AQT scores was .337; the CVT-NC was related to the P/F ($\underline{r}$ = -.274), but AQT scores were not related to the P/F. This finding provides support for selection tests, such as the CVT, that measure abilities to rapidly process information, manipulate the information in STM, make comparisons of this information held in STM with new incoming information, and make decisions based on such comparisons. The CVT is similar in processing demands to the CIP tests of Hunt et al. (21). All subjects tested on the CVT possessed the knowledge requirements to perform the task (i.e., all knew concepts such as, triangle, pentagon, rectangle, and directions N, S, E, W, etc., prior to testing). Nonetheless, subjects differed in their ability to accurately perform the CVT, a task that required many basic processes of the CIP tests of Hunt et al. (21). In the fifth experiment by Hunt et al. (21), a statement (e.g., * ABOVE +) was presented, subjects pressed a button when they comprehended the statement, then a picture of the * and + was presented in various spatial arrangements. Subjects responded true or false. Statement encoding and decision times (in response to the picture) were measured. With negative statements, the encoding and decision times for high verbals were approximately half that of low verbals. For the CVT, higher accuracy was directly related to passing primary flight training and higher AQT scores. The present results indicate that the CVT may be measuring critical basic cognitive processes that mediate both general intelligence test performance and success in naval aviation. Aside from predicting success in naval aviation, the CVT has the potential to measure relevant, basic cognitive processing capabilities without bias against population minorities, assuming that all college graduates possess the necessary background knowledge required to perform the CVT.

mation bits contained in the chunk. Presumably, individuals possessing superior capabilities to encode, maintain, and manipulate a greater amount of information in STM faster, might succeed at a higher rate in flight training. The results of the regression analyses presented above support this assertion.

Basic cognitive processing ability has been related to intelligence. Hunt et al. (21) listed the following basic cognitive processes, collectively referred to as "Current Information Processing" (CIP): processes by which people detect stimuli, locate them in time and space, integrate sequences of stimuli over time to form a stable percept, transform the percept during problem solving, coordinate them with respect to stored representations of highly overlearned codes, and abstract information from the percept and place it in long-term memory. Hunt et al. (21) compared subjects (college freshmen) having high-versus-low verbal ability (measured by the Washington Pre-College Test) and found that high verbals displayed the following superior CIP abilities: ability to make a rapid conversion from a physical representation to a conceptual meaning; ability to maintain information in short-term memory about the order of stimulus presentation; and ability to be more rapid in the manipulation of data in short-term memory. Hunt et al. (21) regarded information processing capacity as a more basic mental ability than the composite skill of verbal aptitude. Performance on verbal aptitude tests is largely determined by what a person knows and not just how rapidly someone can perform basic, simple, cognitive processes such as the CIPs described above.

The correlation between the CVT-NC and the AQT scores was .337; the CVT-NC was related to the P/F ($\underline{r}$ = -.274), but AQT scores were not related to the P/F. This finding provides support for selection tests, such as the CVT, that measure abilities to rapidly process information, manipulate the information in STM, make comparisons of this information held in STM with new incoming information, and make decisions based on such comparisons. The CVT is similar in processing demands to the CIP tests of Hunt et al. (21). All subjects tested on the CVT possessed the knowledge requirements to perform the task (i.e., all knew concepts such as, triangle, pentagon, rectangle, and directions N, S, E, W, etc., prior to testing). Nonetheless, subjects differed in their ability to accurately perform the CVT, a task that required many basic processes of the CIP tests of Hunt et al. (21). In the fifth experiment by Hunt et al. (21), a statement (e.g., * ABOVE +) was presented, subjects pressed a button when they comprehended the statement, then a picture of the * and + was presented in various spatial arrangements. Subjects responded true or false. Statement encoding and decision times (in response to the picture) were measured. With negative statements, the encoding and decision times for high verbals were approximately half that of low verbals. For the CVT, higher accuracy was directly related to passing primary flight training and higher AQT scores. The present results indicate that the CVT may be measuring critical basic cognitive processes that mediate both general intelligence test performance and success in naval aviation. Aside from predicting success in naval aviation, the CVT has the potential to measure relevant, basic cognitive processing capabilities without bias against population minorities, assuming that all college graduates possess the necessary background knowledge required to perform the CVT.

CONCLUSIONS AND RECOMMENDATIONS

Based on data from 451 student naval aviators, the CVT-NC was a consistent predictor of success during Navy primary flight training. This measure of performance accuracy on the CVT predicted to all three criterion measures of Navy primary flight training success: P/F, FG, and CS. With the BI score (a test already administered by the U.S. Navy) and the CVT-NC score, 10.24% of the P/F variance was accounted for. Given that the BI is already available, this increase in prediction capability from one additional test is quite significant. These findings provided a critical external validation for the methodology of Random Sampling of Domain Variances (5), which was developed as part of the AHFETE Program and to which the success of the CVT as a predictor test was, no doubt, largely due. The present analyses provided strong support for using the Random Sampling of Domain Variances methodology in selection research, as well as in human factors engineering design efforts.

The following recommendations are offered:

1. The CVT should be further evaluated in a cross-validation study and then considered for implementation as a performance-based selection test.

2. The method of Random Sampling of Domain Variances should be used to design future selection tests. This methodology provided strong internal validity for incorporating real-world task demands into the "laboratory" test. The analyses presented herein also provided strong external validity in support of this methodology.

3. The data included in this report should be analyzed further--via the hierarchical factor analysis methods (8), also developed under the AHFETE program. These modifications to standard hierarchical factor analytic techniques were developed for the specific task of identifying particular, individual underlying internal processes (7) that could then be related to specific criterion measures. Such follow-on analyses of the present data should allow identification of basic perceptual/cognitive processes critical to naval aviation.

REFERENCES

1. Waldeisen, L.E., Eyraud, M.Y., and Petho, F.C., Augmentation of Human Factors Engineering Technology Efforts Program Plan, Human Performance Sciences Department Technical Memorandum, Naval Aerospace Medical Research Laboratory, Pensacola, FL, February 1981.

2. Navy Research Advisory Committee, Man-machine Technology in the Navy, NRAC 80-8, December 1980.

3. Comptroller General Report to the Congress of the United States, Effectiveness of U.S. Forces can be Increased Through Improved Weapons System Design, PSAD-81-17, 29 January 1981.

4. Doll, R.E., Naval Flight Officer Function Analysis, Vol. 3: P-3C Tactical Coordinator (TACCO), Naval Aerospace Medical Research Laboratory, Pensacola, FL, 1972.

5. Wherry, R.J., Jr., Theoretical Development for Identifying Underlying Internal Processes: Volume 3. Random Sampling of Domain Variances: A New Experimental Methodology, NAMRL/NADC Joint Report, NAMRL SR86-3/NADC 86105-60, Naval Aerospace Medical Research Laboratory, Pensacola, FL, September 1986. (AD A176 350)

6. Wherry, R.J., Jr., Theoretical Development for Identifying Underlying Internal Processes: Volume 4. Task Domains of Naval Flight Officers (NFOs), NAMRL/NADC Joint Report, NAMRL SR86-4/NADC 86105-60, Naval Aerospace Medical Research Laboratory, Pensacola, FL, September 1986. (AD A176 351)

7. Wherry, R.J., Jr., Theoretical Development for Identifying Underlying Internal Processes: Volume 1. The Theory of Underlying Internal Processes, NAMRL/NADC Joint Report, NAMRL SR86-1/NADC 86105-60, Naval Aerospace Medical Research Laboratory, Pensacola, FL, September 1986. (AD A176 466)

8. Wherry, R.J., Jr., Theoretical Development for Identifying Underlying Internal Processes: Volume 2. Modifications to Hierarchical Factor Analysis: Positive Manifold (POSMAN) Rotations, NAMRL/NADC Joint Report, NAMRL SR86-2/NADC 86105-60, Naval Aerospace Medical Research Laboratory, Pensacola, FL, September 1986. (AD A176 349)

9. Morrison, T.R., "Perceptual Processing Capabilities Relevant to Design of Visual Displays for Information Management in Advanced Airborne Weapon Systems." In Proceedings of North Atlantic Treaty Organization Advisory Group for Aerospace Research and Development Aerospace Medical Panel Symposium on "Information Management and Decision Making in Advanced Airborne Weapon Systems," Toronto, Canada, 14-18 April 1986, pp. 10-1 to 10-9.

10. Damos, D.L. and Gibb, G.D., Development of a Computer-based Aviation Selection Test Battery, NAMRL-1319, Naval Aerospace Medical Research Laboratory, Pensacola, FL, August 1986. (AD A179 997)

11. Griffin, G.R., Morrison, T.R., Amerson, T.L., and Hamilton, P.V., Predicting Air Combat Maneuvering (ACM) Performance: Fleet Fighter ACM Readiness Program Grades as Performance Criteria, NAMRL-1333, Naval Aerospace Medical Research Laboratory, Pensacola, FL, October 1987.

12. Statistical Analysis System (SAS) Institute, SAS Version 5.16, Cary, NC.

13. Shoenberger, R.W., Wherry, R.J., Jr., and Berkshire, J.R., Predicting Success in Aviation Training, Report No. 7, Naval School of Aviation Medicine, Pensacola, FL, September 1963.

14. Guilford, J.P., Fundamental Statistics in Psychology and Education, McGraw-Hill Book Company, NY, 1965.

15. Sternberg, S., "High Speed Scanning in Human Memory." Science, Vol. 153, pp. 652-654, 1966.

16. Sternberg, S., "Two Operations in Character Recognition: Some Evidence from RT Measurement." Perception and Psychophysics, Vol. 2, pp. 45-53, 1967.

17. Sternberg, S., "Memory Scanning: New Findings and Current Controversies." Quarterly Journal of Experimental Psychology, Vol. 27, pp. 1-32, 1975.

18. Klatzky, R.L., Human Memory: Structures and Processes, W. H. Freeman and Company, San Francisco, CA, 1980.

19. Miller, G.A., "The Majical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information." Psychological Review, Vol. 63, pp. 81-97, 1956.

20. Glanville, A.D., and Dallenbach, K.M., "The Range of Attention." American Journal of Psychology, Vol. 41, pp. 207-236, 1929.

21. Hunt, E., Lunneborg, C., and Lewis, J., "What Does it Mean to be High Verbal?" Cognitive Psychology, Vol. 7, pp. 194-227, 1975.